DOCUMENT RESUME

ED 320 932 TM 015 151

AUTHOR Green, Suzy A.
TITLE Power Analysis in Repeated Measures Analysis of Variance with Heterogeneously Correlated Trials.
PUB DATE Apr 90
NOTE 41p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Correlation; Data Analysis; Effect Size; *Estimation (Mathematics); *Test Interpretation
IDENTIFIERS Heterogeneity of Variance; *Power (Statistics); *Repeated Measures Design

ABSTRACT

The issue of power in exploratory single group repeated measures analysis of variance is explored under the condition of heterogeneous correlations across the repeated measures trials. Tables developed from 13 real and hypothetical data sets (70 cells) show how well both univariate and multivariate power is estimated under a variety of conditions of non-sphericity and effect sizes. It was shown that a reasonable estimate of power could be obtained based on the mean off-diagonal population correlation and J. Cohen's (1965) estimate of effect size for a one-way analysis of variance. It was recommended that researchers base their estimates of this mean population correlation on what they know about their instrument's test/retest reliability under treatment conditions. One data table and nine figures are included. One appendix lists the sources of the data sets, and another appendix provides a computer program for estimating power. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Suzy A. Green
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED320932

Power Analysis in Repeated Measures Analysis of Variance with Heterogeneously Correlated Trials

Suzy A. Green

Ohio University

Paper presented at the annual meeting of the American Educational Research Association, Boston, April, 1990.

TM015151

BEST COPY AVAILABLE

Abstract

The study explores the issue of power in exploratory single group repeated measures analysis of variance under the condition of heterogeneous correlations across the repeated measures trials. Tables developed from real and hypothetical data sets show how well both univariate and multivariate power is estimated under a variety of conditions of non-sphericity and effect sizes. It was shown that a reasonable estimate of power could be obtained based on the mean off-diagonal population correlation and Cohen's estimate of effect size for a one-way analysis of variance. It was recommended that researchers base their estimate of this mean population correlation on what they know about their instrument's test/retest reliability under treatment conditions.

Power Analysis in Repeated Measures Analysis of Variance with Heterogeneously Correlated Trials

Introduction

The study explores the issue of power in repeated measures analysis of variance under the condition of heterogeneous correlations between the repeated measures trials. Tables developed from real and hypothetical data sets show how well both univariate and multivariate power is estimated under a variety of conditions of non-sphericity and effect sizes. The question of interest is: How well is the repeated measures effect size estimated when a single correlation coefficient is used as a population parameter to represent the correlations among the repeated measures when the correlation matrix is heterogeneous in a single group exploratory design? This question breaks into two main points: (a) If we use a single estimate for ρ, will it provide a good estimate of population power?, and (b) What should be used for the single estimate? The paper continues with background perspectives on the findings, methods, results, the development of the question of interest, conclusions, the importance of the study, and tables supporting the findings.

Background and Perspectives

Power analysis has been studied for several statistical procedures. However, the importance of *a priori* determination of the probability that a statistical test will yield statistically significant results has not been in evidence in the literature (Cohen, 1988). Cohen's text gives many examples of power calculations for various research designs. However, it does not include power computations and tables for repeated measures analysis of variance designs except in its most basic form, the dependent t test. This area has been

explored in part (Barcikowski & Robey, 1985, 1984a, 1984b, and Robey & Barcikowski, 1984), yet the impact on power under the specific condition of heterogeneity of the occasions correlation matrix has not been throughly addressed. This study will attempt to determine whether a single representative estimate of a heterogeneous trials matrix will be adequate for use with current power analysis methods.

Under valid assumptions (sphericity/circularity and more stringently, under uniformity, or constant correlations among the repeated measures trials) the exploratory single-group univariate F test is always more powerful than the multivariate F test, due to the greater number of denominator degrees of freedom in the univariate test when epsilon, ε, is equal to one. Under most circumstances, in an exploratory design, the assumption that sphericity is tenable is unlikely to be upheld. In univariate repeated measures analyses, a routine method to compensate for the violation of the sphericity assumption is the use of an adjusted univariate F test to control the risk of type I error. The adjustment consists of estimating the sphericity parameter, ε, and then multiplying the numerator and denominator degrees of freedom by the estimated ε. The critical F value is then found using the adjusted degrees of freedom. Of the two estimates of ε, $\varepsilon\hat{}$ recommended by Greenhouse and Geisser (1959), and $\varepsilon\tilde{}$ recommended by Huynh and Feldt (1976), $\varepsilon\hat{}$ is the more conservative and therefore is well put to exploratory designs in which the value of ε is probably unknown.

When the sphericity assumption is violated, then the power of the adjusted univariate test varies from being more powerful than the multivariate test to being severely less powerful than the multivariate test (Barcikowski and Robey, 1984a; Jensen, 1982). Therefore, Barcikowski and Robey recommend the use of both tests.

Cohen (1965) describes power as a function of the significance criterion, alpha; sample size, n; and effect size, f, the degree to which a phenomenon is specific and nonzero in the population. Barcikowski & Robey (1985) modified Cohen's effect size categories (small, medium, large), based on an equation involving a single estimate of the intercorrelations among repeated measures trials. Barcikowski and Robey suggest when heterogeneous correlations are found among the occasions in repeated measures designs, that the intraclass correlation coefficient might be used to represent the population trials correlation matrix in the calculation of effect size to estimate power (1985, p. 7).

Stevens defines the intraclass correlation, ρ, as dependence among observations, which increases type I error rate (1986, p. 202). Haggard (1958) defines ρ as the correlation between every pair of members in the group. Others similarly defining intraclass correlation are Pearson (1901), Harris (1913), Fisher (1950), and Winer (1971).

Two other measures of a single estimate to represent the trials correlation matrix are the mean of the off-diagonal of the correlation matrix, and the median off-diagonal. While not the same as an intraclass correlation, central tendency measures are routinely used to represent sets of numbers. Here, these two measures are determined from the correlations between the occasions.

Development

Mean Vectors and Variance-Covariance Matrices

The development of the analysis began with contrived data from Willhoft and Schafer (1989). They combined five patterns of mean vectors, representing models of growth and learning

over time, with variance-covariance matrices with three levels of sphericity: spherical ($\varepsilon\hat{} = 1.00$), nearly spherical ($\varepsilon\hat{} = 0.95$), and moderately spherical ($\varepsilon\hat{} = .75$). Willhoft and Schafer state that one factor affecting both univariate and multivariate power in repeated measures designs is the "interaction between the way in which the . . . repeated-factor means are ranked, that is, the 'shape' of the mean vector, μ, and the way in which the . . . elements of Σ [the variance-covariance matrix] are arranged" (p. 2).

Extending some of the examples of Willhoft and Schafer (excluding the three "no growth" patterns of all zeros in the mean vector), 58 other pattern combinations of mean vectors and variance-covariance matrices from real and hypothetical data sets (Appendix A) were examined. Those sets of raw data were first analyzed using the Repeated Analyzer SAS language program (SAS Institute Inc., 1985) created by Barcikowski and Robey (1990). From this program, mean vectors, μ; variance-covariance matrices, Σ; sample sizes, n; number of trials, k; and the non-singular matrix of (k - 1) by k contrast coefficients, C were catalogued. This information was entered into the SAS language program Rho Effect (Appendix B) based on equations in Barcikowski and Robey (1984a, 1985).

Effect Size

In their paper (1985), Barcikowski and Robey show that when the condition of uniformity[1] is met, Cohen's effect size index, f, may be written in both the univariate (f_u) and multivariate (f_m) form as:

$$(1)\quad f_u = f_m = f/\sqrt{1 - \rho}$$

where ρ is a population correlation, such as the intraclass correlation coefficient. Specifically, the univariate and multivariate effect sizes under uniformity are:

$$(2)\quad f_u = f_m = \sqrt{\frac{\sum_{i=1}^{K} (\mu i - \mu .)^2}{K\,\sigma_m^2\,(1 - \rho)}}$$

This equation is Cohen's (1988, p. 275) effect size for a one-way analysis of variance with K independent groups divided by $\sqrt{1 - \rho}$. For the purpose of repeated measures analysis of variance, K equals the number of trials in a single group design. The numerator of this formula can found as the sum of squares between measures in a one-way analysis of variance. In the denominator, σ_m^2 can be found as the population variance for any given measure. In practice, σ_m^2 and ρ are estimated from norm information or from past research/pilot work.

In an exploratory analysis, ρ is an unknown quantity, and difficult to estimate if homogeneity between the repeated measures trials correlation matrix cannot be assumed. Under the limitation of compound symmetry, formula (2) can be used to calculate effect size which is used to calculate power. Actual power is calculated using σ_{uxm}^2, the interaction of units and measures, or the error, variance, which can be estimated as the mean square within residual in a reliability program analysis (e.g., SPSS^X RELIABILITY). Under the condition of circularity,

$$(3)\quad \sigma_{uxm}^2 = \sigma_m^2\,(1 - \rho)$$

Formula (2) employs the estimate of σ_m^2 $(1 - \rho)$ in the denominator. The effect size calculated may be over- or underestimated using a single estimate of ρ. The answer to the question of interest rests on the magnitude of the over- or underestimation.

Methods

Data Sets

Seventy cells were catalogued from 13 data sets. Six data sets came from published sources, some real, some hypothetical. The remaining data sets were collected from practitioners at Ohio University in Athens, OH, and were all real data. The studies came from a wide variety of disciplines, such as hearing and speech sciences, botany, interpersonal communication, curriculum and instruction, and osteopathy. While most of the practitioners' data sets were exploratory, only two were single group designs. The analyses were performed on each group separately, as if each were a single group. The same procedure was followed when there was more than one measure per occasion. Only one measure was analyzed at a time. The hypothetical data sets were separated the same way.

Procedure

The Rho Effect program was used to calculate the parameters of interest. Using information from the Repeated Analyzer program, among the measures that were calculated are the following: (a) RBAR, the average of the lower triangular half of the trials correlation matrix; (b) RMED, the median of the lower triangular half of the trials correlation matrix; (c) MES_POP, the multivariate effect size in the population; (d) UES_POP, the population univariate effect size;

(e) ES_MEAN,the effect size estimate based on RBAR;
(f) ES_MED, the effect size estimate based on RMED;
(g) MPOP_POW, multivariate population power; (h) UPOP_POW, univariate population power; (i) MEAN_POW, multivariate power estimated from RBAR; (j) MED_POW, multivariate power estimated from RMED; (k)UMEANPOW, univariate power estimated from RBAR; and (l) UMED_POW, univariate power estimated from RMED.

Intraclass correlation coefficients were not included in the final Rho Effect program. Several intraclass correlation coefficient formulas were investigated (Harris, 1913; Fisher, 1950; and Winer, 1971), but the resulting effect sizes tended to be so large, both in absolute terms and in comparison to population values, as to be useless and so were discarded.

As there was a varied mixture of real and hypothetical data sets, the analysis of how well power is estimated was descriptive in nature. The Microsoft Works for Apple Macintosh Systems spreadsheet (Productivity Software Inc., 1988) was used to build the database of the variables. Comparisons were made of population parameters vs. derived estimates for ρ, f, and power. Statpro (Penton Software Inc., 1985) was used for graphics.

Results

Table 1 lists comparisons of ρ, f, and power. The comparisons over the 70 data set cells are: (a) |SUM|, the sum of the absolute value of the differences to detect the magnitude of the deviation; (b) μ, the mean of the absolute value of the deviations; (c) σ, the standard deviation of the absolute value of the deviations; (d) % OVER, the percentage of analyses that overestimated the population parameter; (e) % SAME, the percentage of analyses that exactly estimated the population parameter; and, (f) % UNDER, the percentage of

analyses that underestimated the population parameter.

The first entry in Table 1 compares the ρ (RPOP) of the population (which was calculated from formula 4) with the mean r of the correlation trials matrix. The second entry compares ρ to the median r of the correlation matrix. The comparison of these two entries shows that the mean r is a better estimator of ρ, in that the difference of $|\rho - r|$ has a smaller mean and variance. Both the mean r and the median r tend to overestimate ρ. Figure 1 depicts the range of deviations of RBAR and RMED from RPOP. It is easy to see that RMED deviated from RPOP much more than did RBAR. The range bar plot for the mean r deviation shows it to be the more accurate estimator of rho, with checkpoints at the minimum deviation score, 25th percentile, median, and maximum deviation score.

The third through sixth entries in Table 1 compare univariate and multivariate population effect sizes with estimates of effect size based on mean r and median r. Comparison of the third and fourth entries show that univariate effect size calculated using mean r is a closer estimate of the population parameter than using median r. As in estimating ρ, both estimates tend to overestimate the population effect size. Visual support for these findings are in Figures 2 and 3. These figures contain histogram frequency categories with ten range intervals. The tenth interval on Figure 2 and the eighth interval on Figure 3 contain zero (indicating the closest estimation of a parameter). ES_MEAN (Figure 2) as an estimator falls into this "best fit" interval surrounding zero 74% of the time compared to ES_MED which fits into the interval (Figure 3) 52.8% of the time. Both figures are negatively skewed, indicating that both ES_MEAN and ES_MED tend to overestimate UES_POP.

Entries five and six in Table 1 show first that mean
median r estimate univariate effect size at least 300% l
than they estimate multivariate effect size. In other v
the deviations of |SUM|, μ, and σ, of mean and median ef
size from multivariate population effect size are over t
times greater than the univariate comparisons. However,
within this category, the median r is the better estimat
Again, both tend to overestimate the parameter.

Entry seven compares two population parameters, and
indicates which has more power: univariate or multivari
Forty-three percent of the time, univariate population p
is greater, compared to the 24% of the time that multiva
power is greater.

Entries eight through 13 compare population power
parameters to estimated power. Entry eight vs. nine sho
that univariate power based on mean r deviates from popu
power less than that based on median r. Univariate mean
also over- and underestimates less often than univariate
median power. Figures 4 and 5 show the overestimation i
negatively skewed histograms. In the univariate mode,
UMEANPOW fits into the interval containing zero 90% of t
time (Figure 4) compared to the UMED_POW's 85.7% (Figure
The deviations of the multivariate calculations tend to
larger than the deviations of the univariate calculation
Figures 6 and 7 show that while MEAN_POW and MED_POW ten
more equally over- and underestimate multivariate power,
the frequency graphs show a more "normal" distributional
shape, that neither fits into the interval containing ze
often as the univariate estimators do. That is, MEAN_PO
into the "zero" interval 75.5% of the time (Figure 6) an
MED_POW (Figure 7) fits into the interval 65.7% of the t
In Table 1, the multivariate comparisons (entries 10 and
likewise support the evidence that the mean r is a more

accurate estimator than the median r in deriving power. Deviations of the median r are several times greater than deviations of the mean r. Entries 12 and 13 show that crossovers between univariate and multivariate parameters and estimates fare less will than non-crossover deviations.

Entry 14 compares the multivariate and univariate mean r power estimates. The sample shows univariate mean power as the greater 63% of the time, and the lesser only 1% of the time. This comparison echoes that of entry seven showing the same balance of results as the population values.

Entries 15 and 16 compare the maximum population power per cell (multivariate or univariate, whichever is greater) with univariate mean power and multivariate mean power. UMEANPOW appears to approximate maximum population values more closely. Figures 8 and 9 show that 82.8% of the time UMEANPOW estimates the parameter within the interval containing zero (Figure 8). In Figure 9, UMED_POW fits into that interval only 54% of the time.

Conclusions

Admittedly, a weakness of the analysis is that with 70 cells stemming from 13 data sets, cells are related. However, this exploratory analysis of the question of interest turned up very consistent results, despite the wide variation in the nature of the studies across data sets. The first main point: If we use a single estimate for ρ, will it provide a good estimate of population power?, can be answered affirmatively, based on the small magnitude of many of the deviation scores and other information in Table 1 and the corresponding figures. As to the second main point: What should the estimate be (e.g., mean r, median r)?, in this sample of cells, it was shown that the mean r is the better estimator of

ρ, f, and power, particularly in the univariate mode. A caution in the use of mean r as an estimator is that it tends to overestimate the parameters. However, this single estimate can be used with Barcikowski and Robey's (1985) sample size tables for a single group repeated measures analysis of variance.

Another option to the question of the second main point on the choice of the estimate lies in the type of correlation coefficient used. An applicable correlation coefficient for the purpose of determining a single estimate of the correlations among heterogeneous trials might be test/retest reliability estimates. This type of correlation coefficient is based on testing the same examinees twice with the same test and then correlating the results. Repeated measures designs and test/retest reliabilities share "hazards" such as carry-over and practice effects, changes in subjects' degree of information and/or ability between trials, and time-interval effects (Allen and Yen, 1979).

In practice, a researcher may have *a priori* access to a measure of the test/retest reliability of an instrument. There are two serious cautions to this suggestion, however. First, these "hazards" may result in an over- or underestimation of the true ρ. Second, actual test/retest reliabilities are not subject to treatment effects as might be applied in a repeated measures design. Any type of systemic treatment employed between trials might have the effect of lowering ρ.[2] The reported test/retest reliability might, *a priori*, be viewed, conservatively as an upper bound of ρ, the degree to which rests on the effects of the treatments on the subjects.

Recommendations

When a researcher is trying to determine *a priori* which r to use to estimate power, there are three routes that may assist:

1. Use Barcikowski and Robey (1985, Table 1, p. 7) effect sizes alone. This route is best employed when the researcher has little or no information about the characteristics of the population from which the sample was drawn, such as in an exploratory design. One can conservatively chose a small effect size assuming a small correlation between measures, or try any of eight other effect sizes based on small, medium, and large correlations for use in Barcikowski and Robey's sample size for power tables. When no other information is available *a priori*, the most conservative effect size will afford the largest sample size in order to detect significant differences.

2. Use Formula (1) with Cohen's effect sizes and a reliability estimate for ρ. This route works best when, although the researcher may not know much about the actual characteristics of the population, the researcher is cognizant of issues that may affect the estimate of ρ. Perhaps the researcher has a reliability coefficient from a test reviewer reference, but plans to sample a very small group. Then the researcher would consider the referenced reliability coefficient as an upper bound for that sample. The researcher would also need to assimilate information about administration and sampling error into the determination of an estimate for ρ.

3. Use Formula (2) with norm information for σ_m^2 and ρ and consider different mean patterns. This route requires

knowledge about the population characteristics as well as issues concerning reliabilities. In an exploratory analysis, it would not be usual for the researcher to have this much information prior to conducting the study.

For an exploratory single group repeated measures analysis of variance design, when some knowledge about the instrument's test/retest reliability under treatment conditions can be obtained, the reseacher will have reasonable results estimating power using route #2. For all the routes, a researcher using a single estimate of ρ will also need to make a knowledgeable decision about which degrees of freedom to employ. This study shows that the univariate test using an estimate of ρ is more powerful than the multivariate test. Using the univariate degrees of freedom with the univariate test would give more accuracy in estimating population univariate power. However, an informed researcher may choose to use the multivariate denominator degrees of freedom, i.e., $(N-K+1)$ vs. the univariate $(N-1)(K-1)$, to arrive at a conservative sample size.

Importance of the Study

The development of sample size/power tables in exploratory single group repeated measures analysis of variance when the trials correlation matrix is heterogeneous addresses an issue not throughly researched. In exploratory designs it is most likely the case that correlation matrices are heterogeneous rather than uniform. A method for estimating effect size in single group repeated measures analyses will assist researchers and practitioners in estimating *a priori* the power of their designs.

References

Allen, M. J., & Yen, W. M. (1979). Introduction to measurement theory. Monterey, CA: Brooks/Cole.

Barcikowski, R. S., & Robey, R. R. (1990, April). Interpreting repeated measures data through "special" mean contrasts. Paper presented at the meeting of the American Educational Research Association, Boston, MA.

Barcikowski, R. S., & Robey, R. R. (1985, April). Sample size selection in single group repeated measures analysis. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Barcikowski, R. S., & Robey, R. R. (1984a, April). Exploratory repeated measures analysis. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Barcikowski, R. S., & Robey, R. R. (1984b). Decisions in single group repeated measures analysis: Statistical tests and three computer packages. The American Statistician, 38, 148-150.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New Jersey: Lawrence Erlbaum.

Cohen, J. (1965). Some statistical issues in psychological research. In B. B. Wolman (Ed.), Handbook of clinical psychology (pp. 95-121). New York: McGraw-Hill.

Fisher, R. A. (1950). Statistical methods for research workers (11th ed.). London: Oliver & Boyd.

Greenhouse, S. W., & Geisser, S. (1959). On methods in analysis of profile data. Psychometrika, 24, 95-112.

Haggard, E. A. (1958). Intraclass correlation and the analysis of variance. New York: Dryden Press.

Harris, J. A. (1913). On the calculation of the intraclass and interclass coefficients of correlation from class moments when the number of possible combinations is large. Biometrika, 9, 446-472.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. Journal of Educational Statistics, 1, 69-82.

Jensen, D. R. (1982). Efficiency and robustness in the use of repeated measurements. Biometrics, 38, 813-825.

Penton Software Inc. (1985). Statpro (Version 2.0 ed.). New York: Penton Software, Inc.

Pearson, K. (1901). Mathematical contributions to the theory of evaluation. X. Supplement to a memoir on skew variation. Philosophical Transactions of the Royal Society of London, Series A, 285-379.

Productivity Software Inc. (1988). Microsoft Works (Version 2.0 ed.). Redmond, WA: Microsoft Corporation.

Robey, R. R., & Barcikowski, R. S. (1984). Calculating the statistical power of the univariate and the multivariate repeated measures analyses for the single group case under various conditions. Educational and Psychological Measurement, 44, 137-143.

SAS Institute Inc. (1985). SAS user's guide: Statistics (Version 5 ed.). Cary, NC: SAS Institute Inc.

Stevens, J. P. (1986). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Lawrence Erlbaum.

Willhoft, J. L, & Schafer, W. D. (1989, March). The effects of mean pattern and non-sphericity on univariate and multivariate analyses for a two-factor repeated measures design. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill.

Footnotes

[1]Using a single estimate of ρ is tantamount to assuming uniformity. However, this does not change the basic question of whether or not a reasonable estimate of power can be found using a single representation of a heterogeneous correlation trials matrix.

[2]However, if the treatment has a linear effect between measures, i.e., no interaction betweeen treatments and subjects, then the use of the reliability as an estimate seems reasonable.

Table 1

Summary Statistics for Selected Differences Scores

VARIABLE	|SUM|	μ	σ	% OVER	% SAME	% UNDER
1) RPOP-RBAR*	2.4396	.0348	.0610	73	10	17
2) RPOP-RMED*	3.8711	.0553	.0681	70	06	24
3) UESPOP-ESMEAN*	3.1506	.0450	.0968	71	17	12
4) UESPOP-ESMED*	3.8689	.0552	.0777	70	06	24
5) MESPOP-ESMEAN	12.9866	.1855	.2153	61	06	33
6) MESPOP-ESMED	12.517	.1788	.1991	61	06	33
7) MPOPPOW-UPOPPOW	5.4239	.0774	.1338	43	33	24
8) UPOPPOW-UMEANPOW*	1.3801	.0197	.0650	37	56	07
9) UPOPPOW-UMEDPOW*	2.0225	.0288	.0865	44	44	12
10) MPOPPOW-MEANPOW*	4.6416	.0663	.1283	38.5	38.5	23
11) MPOPPOW-MEDPOW*	22.0883	.3155	2.1465	38.5	38.5	23
12) UPOPPOW-MEANPOW	29.0947	.4156	2.079	8.5	38.5	53
13) MPOPPOW-UMEANPOW	5.6543	.0807	.1307	49	34	17
14) MEANPOW-UMEANPOW	3.0632	.0437	.0575	63	36	01
15) MAXPOPPOW-UMEANPOW*	2.5143	.0359	.1082	23	56	21
16) MAXPOPPOW-MEANPOW*	4.4659	.0637	.1178	09	37	54

* See figures for more detail.

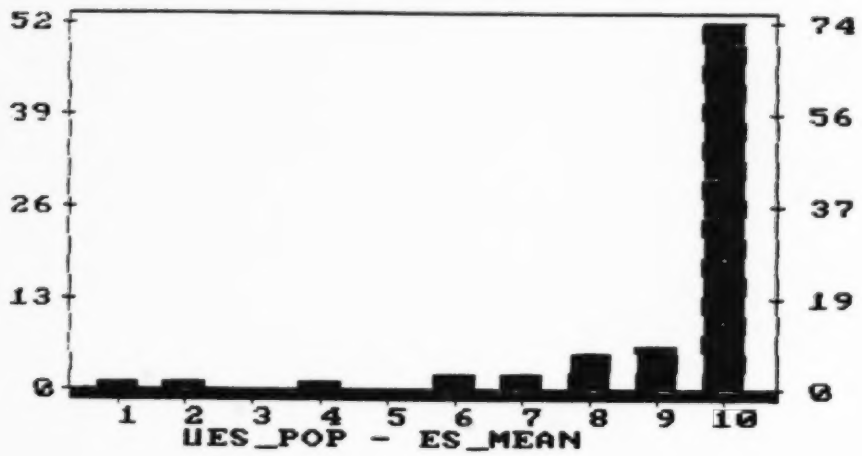
FIGURE 1

RPOP-RBAR vs. RPOP-RMED

Field	Number	Range
rpop-rbar	70	0.32280
rpop-rmed	70	0.49330

Field	Minimum	25th Percentile	Median	75th Percentile	Maximum
rpop-rbar	-0.29110	-0.03870	-8.200000e-3	0.00000	0.03170
rpop-rmed	-0.33630	-0.07345	-8.950000e-3	0.00000	0.15700

FIGURE 2

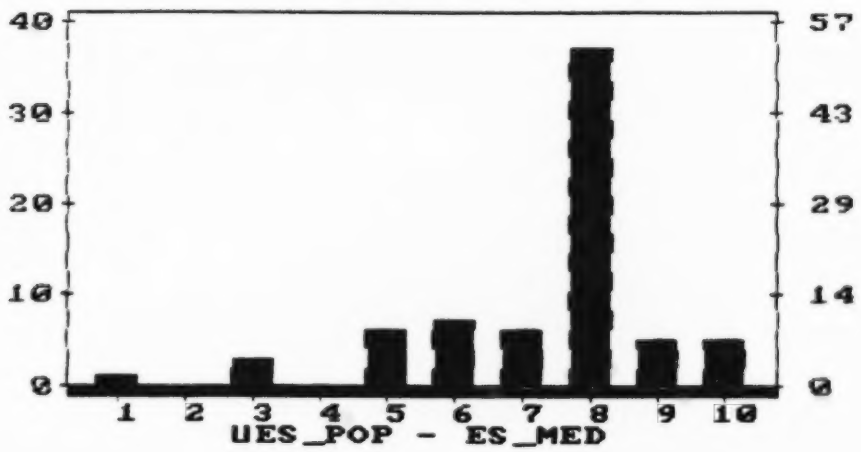


Frequency Histogram

Histogram Frequency Results for Field #25:13 - 14

Bar #	Low	High	Number	Percent
1	-0.5254	-0.4714	1.0000	1.4286
2	-0.4714	-0.4174	1.0000	1.4286
3	-0.4174	-0.3634	0.0000	0.0000
4	-0.3634	-0.3094	1.0000	1.4286
5	-0.3094	-0.2554	0.0000	0.0000
6	-0.2554	-0.2014	2.0000	2.8571
7	-0.2014	-0.1474	2.0000	2.8571
8	-0.1474	-0.0934	5.0000	7.1429
9	-0.0934	-0.0394	6.0000	8.5714
10	-0.0394	0.0146	52.0000	74.2857

FIGURE 3

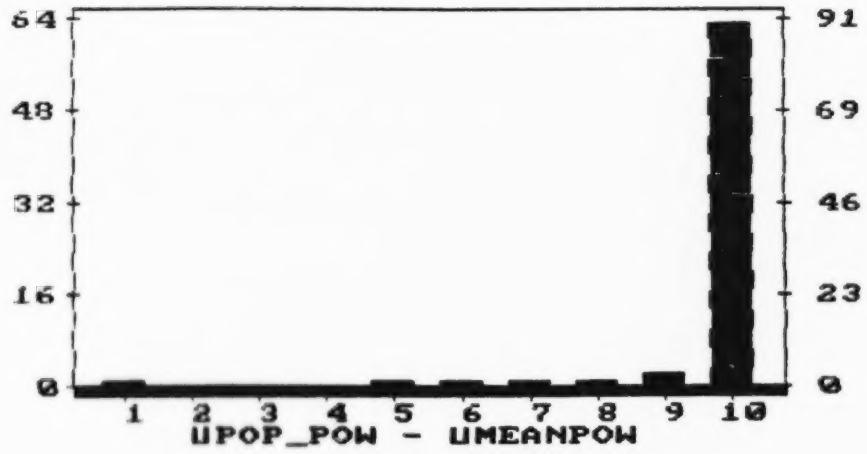


Frequency Histogram

Histogram Frequency Results for Field #26:13 - 15

Bar #	Low	High	Number	Percent
1	-0.3841	-0.3345	1.0000	1.4286
2	-0.3345	-0.2850	0.0000	0.0000
3	-0.2850	-0.2354	3.0000	4.2857
4	-0.2354	-0.1859	0.0000	0.0000
5	-0.1859	-0.1363	6.0000	8.5714
6	-0.1363	-0.0868	7.0000	10.0000
7	-0.0868	-0.0373	6.0000	8.5714
8	-0.0373	0.0123	37.0000	52.8571
9	0.0123	0.0618	5.0000	7.1429
10	0.0618	0.1114	5.0000	7.1429

FIGURE 4

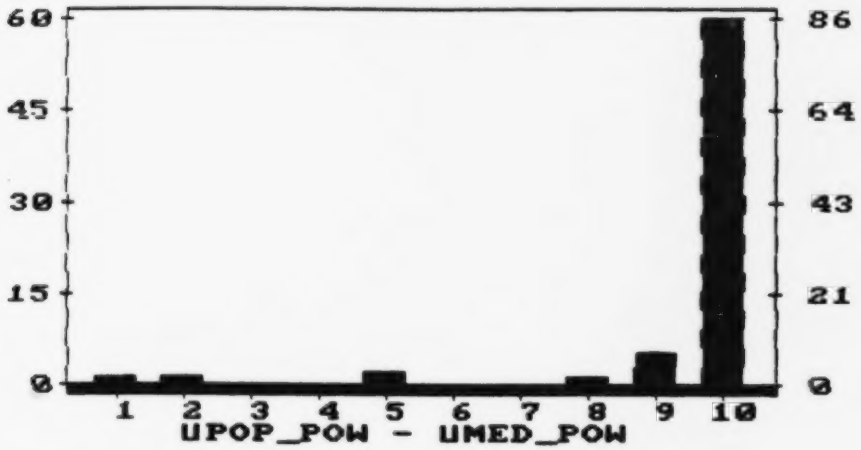


Frequency Histogram

Histogram Frequency Results for Field #20:9 - 10

Bar #	Low	High	Number	Percent
1	-0.4330	-0.3895	1.0000	1.4286
2	-0.3895	-0.3459	0.0000	0.0000
3	-0.3459	-0.3024	0.0000	0.0000
4	-0.3024	-0.2588	0.0000	0.0000
5	-0.2588	-0.2153	1.0000	1.4286
6	-0.2153	-0.1718	1.0000	1.4286
7	-0.1718	-0.1282	1.0000	1.4286
8	-0.1282	-0.0847	1.0000	1.4286
9	-0.0847	-0.0411	2.0000	2.8571
10	-0.0411	2.434600e-3	63.0000	90.0000

FIGURE 5

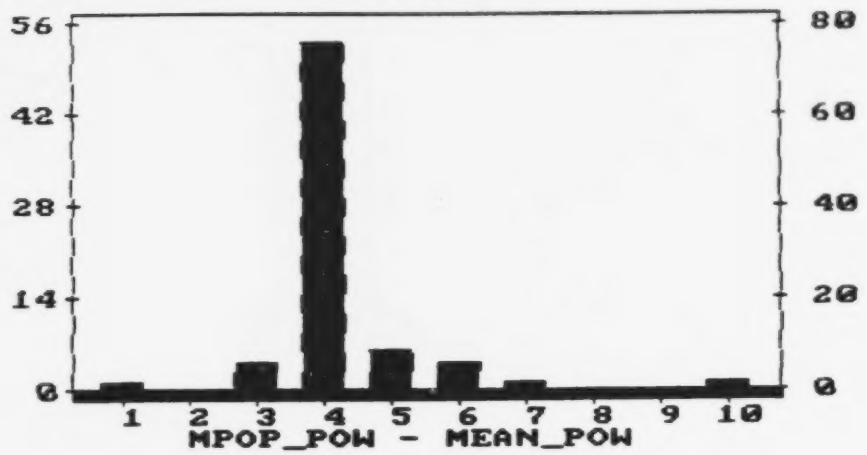


Frequency Histogram

Histogram Frequency Results for Field #21:9 - 11

Bar #	Low	High	Number	Percent
1	-0.4953	-0.4432	1.0000	1.4286
2	-0.4432	-0.3910	1.0000	1.4286
3	-0.3910	-0.3388	0.0000	0.0000
4	-0.3388	-0.2867	0.0000	0.0000
5	-0.2867	-0.2345	2.0000	2.8571
6	-0.2345	-0.1823	0.0000	0.0000
7	-0.1823	-0.1302	0.0000	0.0000
8	-0.1302	-0.0780	1.0000	1.4286
9	-0.0780	-0.0258	5.0000	7.1429
10	-0.0258	0.0263	60.0000	85.7143

FIGURE 6

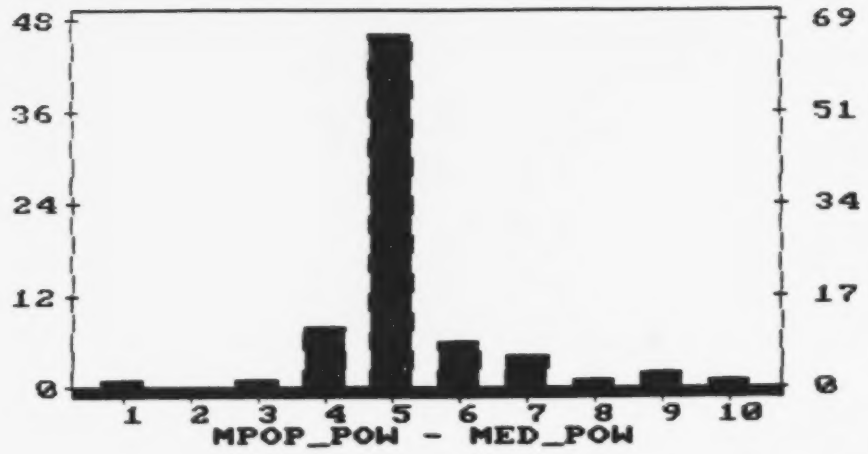


Frequency Histogram

Histogram Frequency Results for Field #16:6 - 7

Bar #	Low	High	Number	Percent
1	-0.4948	-0.3635	1.0000	1.4286
2	-0.3635	-0.2322	0.0000	0.0000
3	-0.2322	-0.1009	4.0000	5.7143
4	-0.1009	0.0304	53.0000	75.7143
5	0.0304	0.1617	6.0000	8.5714
6	0.1617	0.2931	4.0000	5.7143
7	0.2931	0.4244	1.0000	1.4286
8	0.4244	0.5557	0.0000	0.0000
9	0.5557	0.6870	0.0000	0.0000
10	0.6870	0.8183	1.0000	1.4286

FIGURE 7

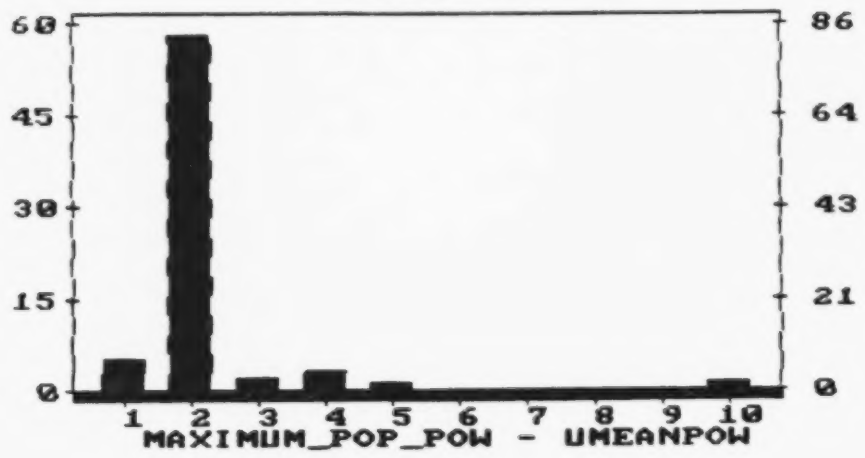


Frequency Histogram

Histogram Frequency Results for Field #17:6 - 8

Bar #	Low	High	Number	Percent
1	-0.4187	-0.3324	1.0000	1.4286
2	-0.3324	-0.2461	0.0000	0.0000
3	-0.2461	-0.1598	1.0000	1.4286
4	-0.1598	-0.0735	8.0000	11.4286
5	-0.0735	0.0129	46.0000	65.7143
6	0.0129	0.0992	6.0000	8.5714
7	0.0992	0.1855	4.0000	5.7143
8	0.1855	0.2718	1.0000	1.4286
9	0.2718	0.3581	2.0000	2.8571
10	0.3581	0.4444	1.0000	1.4286

FIGURE 8

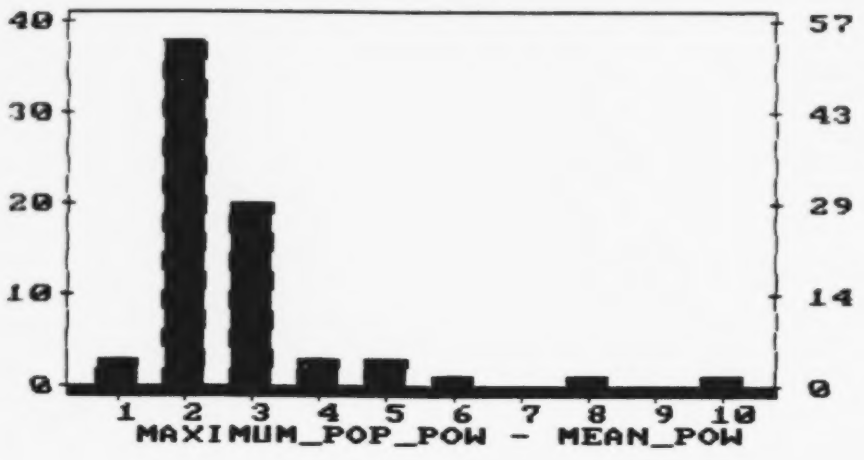


Frequency Histogram

Histogram Frequency Results for Field #28:max-umean

Bar #	Low	High	Number	Percent
1	-0.1884	-0.0887	5.0000	7.1429
2	-0.0887	0.0109	58.0000	82.8571
3	0.0109	0.1106	2.0000	2.8571
4	0.1106	0.2103	3.0000	4.2857
5	0.2103	0.3100	1.0000	1.4286
6	0.3100	0.4096	0.0000	0.0000
7	0.4096	0.5093	0.0000	0.0000
8	0.5093	0.6090	0.0000	0.0000
9	0.6090	0.7086	0.0000	0.0000
10	0.7086	0.8083	1.0000	1.4286

FIGURE 9



Frequency Histogram

Histogram Frequency Results for Field #27:max-mmean

Bar #	Low	High	Number	Percent
1	-0.1988	-0.0971	3.0000	4.2857
2	-0.0971	4.551120e-3	38.0000	54.2857
3	4.551120e-3	0.1062	20.0000	28.5714
4	0.1062	0.2079	3.0000	4.2857
5	0.2079	0.3096	3.0000	4.2857
6	0.3096	0.4113	1.0000	1.4286
7	0.4113	0.5130	0.0000	0.0000
8	0.5130	0.6146	1.0000	1.4286
9	0.6146	0.7163	0.0000	0.0000
10	0.7163	0.8180	1.0000	1.4286

Appendix A

Data Sets

Davidson, M. L. (1972). Univariate versus multivariate tests in repeated-measurements experiments. Psychological Bulletin, 77, pp. 446-452.

Fucci, D., Petrosino, L., Harris, D., Randolph-Tyler, E., & Wagner, S. (1989). Lingual vibrotactile threshold shift during magnitude-estimation responses and scaling behavior. Perception and Psychophysics, 46, pp. 275-279.

Looney, S. W. & Stanley, W. B. (1989). Exploratory repeated measures analysis of variance for two or more groups. The American Statistician, 43, pp. 220-225.

Nellessen, J. E. (1989). Population differentiation in *andropogon virginicus* L. between abandoned coal strip mine spoil and old field habitats in Ohio. Unpublished doctoral dissertation, Ohio University, Athens, OH.

Papa, M. J., & Graham, E. E. (1989). Gender and function-oriented behavior in small groups: An examination of problem-solving processes and outcomes. Manuscript submitted for publication.

Rich, C. E. (1983). Repeated measures designs. In R. S. Barcikowski (Ed.), Computer packages and research designs, with annotations of input and output from the BMDP, SAS, SPSS, and SPSS^{X} statistical packages, Vol. 1 BMDP, Vol. 2 SAS, Vol. 3 SPSS^{X}. Lanham, MD: University Press of America.

Reven, L. (1989). Cognitive capacity usage as it relates to a subject's functional reading levels. Unpublished doctoral dissertation, Ohio University, Athens, OH.

Scuitto, M. (1988). Efficacy of a range of motion warm water exercise program for women with osteoporosis. Paper presented at the meeting of the Ohio University College of Osteopathic Medicine, Athens, OH.

Taylor, R. Y. (1988). An examination of comprehension competence in narrative and expository material using an informal reading inventory with probed and free recall questioning. Unpublished doctoral dissertation, Ohio University, Athens, OH.

Timm, N. H. (1975). Multivariate analysis with applications in education and psychology. Monterey, CA: Brooks/Cole, pp. 232-233, 509-510.

Walkowski, S. (1988). The effect of examination stress on the cellular immunocompetencies of medical students and their spouses. Paper presented at the meeting of the Ohio University College of Osteopathic Medicine, Athens, OH.

Willhoft, J. L., & Schafer, W. D. (1989). The effects of mean pattern and non-sphericity on univariate and multivariate analyses for a two-factor repeated measures design. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Appendix B

Rho Effect

```
*
THE FOLLOWING PROGRAM WAS WRITTEN BY ROBERT S. BARCIKOWSKI AND
SUZY A. GREEN DURING FALL 1989 AND WINTER 1990.  IT IS BASED ON
EQUATIONS IN BARCIKOWSKI AND ROBEY (1984, 1985);
*
*
*
*
*
;
DATA FISHER;
*  FISHER IS AN ITERATIVE SEQUENCE OF STATEMENTS THAT RETURNS
   THE F VALUE REQUIRED FOR SIGNIFICANCE GIVEN, D1, D2, AND
   ALPHA.
;
ALPHA = .05;  * THE LEVEL OF SIGNIFICANCE;
N = 25;  * NUMBER OF UNITS;
K = 5;  * NUMBER OF MEASURES;
F = 1; * STARTING NUMBER FOR F;
D1 = K-1;  * NUMERATOR DEGREES OF FREEDOM;
UD2 = (N-1)*(K-1);  * UNIVARIATE DEGREES OF FREEDOM;
MD2 = N-K+1;  * MULTIVARIATE DEGREES OF FREEDOM
;
* THE FOLLOWING LOOP DETERMINES THE F VALUES
;
      DO J = 1 TO 2;
      IF (J EQ 1) THEN D2 = UD2;
      ELSE D2 = MD2;
      PDIF = 1.00;
      DO I = 1 BY 1;
      P = 1 - PROBF(F,D1,D2);
```

```
      DIF = ABS(P - ALPHA);
      IF (DIF GT PDIF) THEN GO TO A;
      PDIF = DIF;
      F = F + .001;
      END;
      A: F = F - .001;
      IF (J EQ 1) THEN UF = F;
      ELSE MF = F;
      END;
KEEP MF UF ALPHA N K;
OUTPUT;
PROC PRINT;
PROC MATRIX PRINT FUZZ;
FETCH FS DATA = FISHER;
UF = FS(1,4);
MF = FS(1,5);
ALPHA = FS(1,1);
N = FS(1,2);
K = FS(1,3);
U = -.3108 /
    -.2004 /
    -.0524 /
    -.0564 /
    -.0536;  * GENERATED DATA, 5 TRIALS, ONE GROUP,
               COLUMN VECTOR OF THE REP. MEAS. SAMPLE MEANS
;
MK = (K * K - K) #/2;  * THE NUMBER OF CORRELATIONS TO BE PLACED
                         IN TRIAG
;
* WE MUST FIRST PLACE ELEMENTS IN TRIAG BEFORE IT CAN APPEAR ON
  THE LEFT HAND SIDE OF AN = SIGN
;
TRIAG = J(MK,1,1);  * PLACE MK ONE'S IN TRIAG
;
```

```
SIGMA = 1.2852 1.0553  .9928 1.0284  .682  /
        1.0553 1.1904 1.0677 1.098   .6788 /
         .9928 1.0677 1.0922 1.0315  .6786 /
        1.0284 1.098  1.0315 1.1795  .6921 /
         .682   .6788  .6786  .6921  .7122 ;
VAR = DIAG(SIGMA);
STD = SQRT(VAR);
STDINV = INV(STD);
COR = STDINV * SIGMA * STDINV;
*  PLACE THE LOWER TRIANGULAR HALF OF COR INTO A VECTOR (TRIAG)
;
CNT = 0;
DO I = 1 TO K;
DO J = 1 TO I;
IF (I EQ J) THEN GO TO B;
CNT = CNT + 1;
TRIAG(CNT,1) = COR(I,J);
B: END;
END;
*  RANK ORDER THE CORRELATIONS IN TRIAG
;
R_TRIAG = TRIAG;
ATRIAG = RANK(TRIAG);
TRIAG(RANK(TRIAG),) = R_TRIAG;
OTRIAG = TRIAG
;
*  FIND THE MEDIAN OF THE VALUES IN TRIAG
;
MID = INT(MK #/ 2);
MID1 = MID + 1;
RMED = (TRIAG(MID,1) + TRIAG(MID1,1)) #/ 2;
IF MOD(MK,2) EQ 0 THEN RMED = RMED;
  ELSE RMED = TRIAG(MID1,1);  * THE MEDIAN CORRELATION
;
*  FIND THE AVERAGE CORRELATION IN TRIAG
```

```
;
RBAR = SUM(TRIAG) #/ MK;  * THE AVERAGE CORRELATION;
C = 1 0 0 0 -1 /
    0 1 0 0 -1 /
    0 0 1 0 -1 /
    0 0 0 1 -1;
* NONSINGULAR MATRIX OF K - 1 BY K CONTRAST COEFFICIENTS
;
* REDEFINE THE TRANSFORMATION MATRIX C BY ORTHONORMALIZING THE
  COEFFICIENTS, GRAM-SCHMIDT ORTHONORMALIZATION
;
C = C';
GS CTEMP T LINDEP C;
C = CTEMP';  * ROW ORTHONORMALIZED CONTRAST COEFFICIENTS
;
SIG = C * SIGMA * C';
        EIGEN EVALS EVECS SIG;
        DETERM = DET(SIG);
C = EVECS' * C;  * THE RESULTANT C CONTAINS ORTHONORMALIZED
                   CONTRAST COEFFICIENTS SUCH THAT C * SIGMA *
                   C' IS A DIAGONAL MATRIX
;
;
CONTRAST = C * U;  * ORTHONORMALIZED CONTRASTS ON THE REPEATED
                     MEASURES MEANS
;
CON2 = CONTRAST ##2;  * SQUARED CONTRASTS, NEEDED FOR EFFECT
                        SIZES
;
SUM_CON2 = SUM(CON2);  * SUM OF THE SQUARED CONTRASTS
;
CON_VAR = C * SIGMA * C';  * A DIAGONAL MATRIX WITH THE CONTRAST
                             VARIANCES ON THE DIAGONAL
;
```

```
CON_VAR2 = CON_VAR * CON_VAR;  * DIAGONAL MATRIX WITH SQUARED
                                 VARIANCES ON THE DIAGONAL, USED
                                 TO CALCULATE GGI
;
SSB = N * (U' * C' * C * U);  * SUM OF SQUARES BETWEEN MEASURES
;
SSE = TRACE(CON_VAR);  * THE SUM OF THE CONTRAST VARIANCES IS
                         THE SUM OF SQUARES ERROR
;
SSE2 = TRACE(CON_VAR2);  * SUM OF THE SQUARED VARIANCES, USED TO
                           CALCULATE GGI
;
VAR_UXM = SSE #/ (K-1);  * THE INTERACTION VARIANCE, ALSO THE
                           ERROR VARIANCE (EQ 3, 84)
;
VAR_M = TRACE(SIGMA) #/K; * THE ESTIMATE (POOLED ACROSS
                            MEASURES) OF THE VARIANCE OF A
                            SINGLE MEASURE
;
GGI = (SSE * SSE) #/ ((K-1) * SSE2);  * THE GREENHOUSE GEISSER
                                        ESTIMATE OF SPHERICITY
                                        (EPSILON, ER 1, 84)
;
SSWRES = SSE * (N-1);  * SUM OF SQUARES WITHIN PEOPLE RESIDUAL
                         AS FOUND IN RELIABILITY PROGRAM
;
SSWP = SSB + SSWRES;  * SUM OF SQUARES WITHIN PEOPLE AS FOUND IN
                        RELIABILITY PROGRAM
;
SSTOT = SSB + (VAR_M * ((N * K)-K));  * SUM OF SQUARES TOTAL
;
SSBP = SSTOT - SSWP;  * SUM OF SQUARES BETWEEN PEOPLE AS FOUND
                        IN RELIABILITY PROGRAM
;
KINV = 1 #/ K;  * ONE OVER K
;
```

```
INV_CV = INV(CON_VAR);  * DIAGONAL MATRIX WITH ONE OVER THE
                          CONTRAST VARIANCES ON THE DIAGONAL
;
RATIO = INV_CV * CON2;  * VECTOR CONTAINING THE RATIOS OF THE
                          CONTRASTS SQUARED TO THE CONTRAST
                          VARIANCES
;
MES_POP = SQRT (KINV * SUM(RATIO));  * MULTIVARIATE EFFECT SIZE,
                                       CONSIDERED AS THE GOLD
                                       STANDARD (EQ 7, 85)
;
UES_POP = SQRT(((K-1) * SUM_CON2) #/ (K * SSE));  * UNIVARIATE
                                                    EFFECT SIZE
                                                   GOLD STANDARD
                                                    (EQ 6, 85)
;
MDELTA = N * K * MES_POP * MES_POP;  * MULTIVARIATE
                                       NONCENTRALITY PARAMETER,
                                       GOLD STANDARD
                                        (EQ 4, 85)
;
UDELTA = N * K * UES_POP * UES_POP;  * UNIVARIATE NONCENTRALITY
                                       PARAMETER, GOLD STANDARD
                                        (EQ 4, 85)
;
ONE_RBAR = 1 - RBAR;  * ONE MINUS THE MEAN CORRELATION
;
ONE_RMED = 1 - RMED;  * ONE MINUS THE MEDIAN CORRELATION
;
;
ES_MEAN = SQRT ((SSB #/ N) #/ (K * VAR_M * ONE_RBAR));  * EFFECT
                                                          SIZE
                                                          BASED
                                                          ON THE
                                                          MEAN
                                                       CORRELATION
                                                         (EQ11,85)
```

```
;
ES_MED = SQRT ((SSB #/ N) #/ (K * VAR_M * ONE_RMED));  * EFFECT
                                                          SIZE
                                                          BASED
                                                          ON THE
                                                          MEDIAN
                                                       CORRELATION
                                                        (EQ 11,85)
;
MEAN_DEL = N * K * ES_MEAN * ES_MEAN;  * ASSUMING UNIFORMITY,
                                         THE NONCENTRALITY
                                         PARAMETER BASED ON THE
                                         AVERAGE CORRELATION
                                         (EQ 4, 85)
;
MED_DEL = N * K * ES_MED * ES_MED;  * ASSUMING UNIFORMITY, THE
                                      NONCENTRALITY PARAMETER
                                      BASED ON THE MEDIAN
                                      CORRELATION
                                      (EQ 4, 85)
;
DF1 = K-1;  * BOTH UNIVARIATE AND MULTIVARIATE NUMERATOR DEGREES
              OF FREEDOM
;
UDF2 = (K-1) * (N-1);  * UNIVARIATE DENOMINATOR DEGREES OF
                         FREEDOM
;
MDF2 = N - K + 1;  * MULTIVARIATE DENOMINATOR DEGREES OF FREEDOM
;
*-------------------------------------------------------------
;
*    POWER FOR ALL EFFECT SIZES BASED ON THE
     MULTIVARIATE DEGREES OF FREEDOM
;
```

```
D1 = DF1;
D2 = MDF2;
F = MF;
UT = DF1;
V = MDF2;
RLAM = MDELTA;
LINK SEVZEL;
MPOP_POW = PROBNORM(POW1);   * MULTIVARIATE POWER FOR THE GO
                               STANDARD
;
RLAM = MEAN_DEL;
LINK SEVZEL;
MEAN_POW = PROBNORM(POW1);   * POWER BASED ON THE AVERAGE
                               CORRELATION
;
RLAM = MED_DEL;
LINK SEVZEL;
MED_POW = PROBNORM(POW1);    * POWER BASED ON THE MEDIAN
                               CORRELATION
;
*-------------------------------------------------------------
;
*    POWER FOR ALL EFFECT SIZES BASED ON THE
     UNIVARIATE DEGREES OF FREEDOM
;
D1 = DF1;
D2 = UDF2;
F = UF;
UT = DF1;
V = UDF2;
RLAM = UDELTA;
LINK SEVZEL;
UPOP_POW = PROBNORM(POW1);   * UNIVARIATE POWER FOR THE GOLD
                               STANDARD
;
```

```
RLAM = MEAN_DEL;
LINK SEVZEL;
UMEANPOW = PROBNORM(POW1);   * POWER BASED ON THE AVERAGE
                               CORRELATION
;
RLAM = MED_DEL;
LINK SEVZEL;
UMED_POW = PROBNORM(POW1);   * POWER BASED ON THE MEDIAN
                               CORRELATION
;
STOP;
*-------------------------------------------------------------
;
SEVZEL:
*  UT----DEGREES OF FREEDOM FOR THE NUMERATOR OF THE F-TEST
   V-----DEGREES OF FREEDOM FOR THE DENOMINATOR OF THE F-TEST
   RLAM--NONCENTRALITY PARAMETER
   F-----F VALUE REQUIRED FOR SIGNIFICANCE WITH UT AND V
   POW1--SCORE RETURNED TO MAIN PROGRAM WHICH WHEN SENT TO
   PROBNORM DETERMINES THE POWER
;
      TEMP1 = SQRT(2*(UT+RLAM)-(UT+2*RLAM) #/ (UT+RLAM));
      TEMP2 = SQRT(((2*V-1)*UT*F)#/V);
      TEMP3 = SQRT((UT*F)#/V+(UT+2*RLAM)#/(UT+RLAM));
      POW1 = (TEMP1-TEMP2)#/TEMP3;
      RETURN;
```